Hoggan, E. and Brewster, S.A. (2006) Crossmodal spatial location: initial experiments. In*, 4th Nordic Conference on Human-Computer Interaction, 14-18th October 2006.* ACM International Conference Proceeding Series Vol 189, pages pp. 469-472, Oslo, Norway.

Crossmodal Spatial Location: Initial Experiments

Eve Hoggan and Stephen Brewster
Glasgow Interactive Systems Group
University of Glasgow, Glasgow G12 8QQ, UK
{eve, stephen}@dcs.gla.ac.uk

ABSTRACT

This paper describes an alternative form of interaction for mobile devices using crossmodal output. The aim of our work is to investigate the equivalence of audio and tactile displays so that the same messages can be presented in one form or another. Initial experiments show that spatial location can be perceived as equivalent in both the auditory and tactile modalities Results show that participants are able to map presented 3D audio positions to tactile body positions on the waist most effectively when mobile and that there are significantly more errors made when using the ankle or wrist. This paper compares the results from both a static and mobile experiment on crossmodal spatial location and outlines the most effective ways to use this crossmodal output in a mobile context.

Author Keywords

Tactons, Earcons, spatial location, mobility

ACM Classification Keywords

H5.2. User Interfaces: Auditory (non-speech) Feedback, Haptic I/O, Interaction Styles.

INTRODUCTION

The amount of information handled by mobile devices is increasing yet most mobile interfaces have problems displaying such a vast quantity of data due to their reliance on small visual displays. If the information available from these devices is to be accessible to all users in all mobile situations, it will be necessary to address the mobility restrictions enforced by present-day interface designs. In this paper crossmodal interaction using varying 3D audio spatial locations and tactile body locations is discussed as a potential alternative to current output techniques.

Despite the increasing amount of information and functionality available from mobile devices, there is yet to be an improvement on their restricted input/output capabilities. Most mobile interface designs today rely on concepts used in desktop computing with graphical user interfaces communicating information visually through extremely small text and images. Whilst attempting to interact with these small displays, the user's visual focus is shifted away from their primary task. In many mobile situations, such as walking, the user's eyes may be occupied although they are otherwise able to manage information from the mobile device by using their other senses.

It can be unnatural to be forced to interact with the environment around us using only our vision. To take an example, in a dark environment we may choose to use touch for navigation as opposed to vision. Similarly, people with sensory disabilities are often forced to use alternative senses. A common technique used to help the sensory impaired is sensory substitution where one sensory modality is used to supply environmental information normally gathered by another sense. The replacement of a sense by another one could be employed by mobile devices too. It could be argued that mobile device users are 'situationally impaired' for example wearing gloves, being in a very noisy environment, or driving. Using notions drawn from sensory substitution, mobile device users could simply use the appropriate modalities as desired.

The manufacturers of mobile devices like PDAs and phones commonly include audio and vibrotactile feedback in their products. This research will build on these features by using auditory/tactile crossmodal output as a form of sensory substitution which will provide the user with multiple ways of accessing the same information on a mobile device, beyond the traditional GUI and small screen.

CROSSMODAL INTERACTION

If we wish to provide effective crossmodal output to mobile displays, it is first necessary to investigate the different parameters available for manipulation in the auditory and tactile modalities because sensory substitution requires the same information to be encoded and presented interchangeably via both modalities.

There is a significant amount of research on individual modalities. Earcons are a type of non-speech auditory display, which Blattner defines as "non-verbal audio messages that are used in the computer/user interface to provide information to the user" [1]. Similarly, in the tactile domain, Brewster and Brown have developed Tactons [2] for structured vibrotactile messages which can be used to communicate information non-visually.

Crossmodal icons [3] are abstract icons which can be instantiated as either an Earcon or Tacton, such that the resultant Earcons or Tactons are equivalent and can be compared as such. Crossmodal icons allow the same information to be accessible interchangeably via the two different modalities. Initial research into crossmodal icons has shown parameters such as rhythm and texture to be easy to map between the audio and tactile domains. There is, however, no complete set of parameters and our research is focused on identifying what works well across the two modalities.

Spatial location

Any attribute that can specify similar information across modalities is considered to be *amodal* in nature [4]. Like intensity, rate, and rhythmic structure, spatial location is a common type of amodal attribute with great potential for intersensory exploration. Auditory and tactile stimuli can be combined in a number of interesting ways to transform the observer's sense of space on and around the body.

Using a set of Earcons and Tactons as crossmodal icons, the work presented here investigates the use of spatial location as another parameter for use in mobile crossmodal auditory and tactile displays. By investigating ways to map from a tactile location on the body to an audio location in a soundscape and *vice versa*, navigational cues for instance could be presented via Earcons and Tactons, as in Figure1.

Different spatial locations can be created in the audio domain by using 3D sound, as this allows sound sources to be positioned in space around a listener. 3D audio systems can create the effect of sound sources behind, above and below the listener.

Figure 1. 'turn right' cue indicated by audio panned to the right (Earcon) and activation of tactor on right hand side of waist (Tacton).

Different spatial locations can be used in the tactile domain by placing transducers at different locations on the body. The spatial location of transducers has been used successfully by many researchers [5, 6]. To use it as a parameter in crossmodal interaction, it is important to choose the body locations carefully. Cholewiak [5] reports that early work on tactile perception suggested that tactile localization is most precise when the stimulus is close to an anatomical reference point. Ergonomic issues should also be considered so that users feel comfortable with actuators on these particular parts of their body.

The experiments described here have investigated whether spatial locations in the audio and tactile domain can be perceived as equivalent when static or mobile and therefore be used in crossmodal interaction with mobile devices.

EXPERIMENT

The first experiment was conducted to determine which body locations can be mapped most effectively to locations in a 3D audio soundscape whilst the user is stationary.

The most recent iteration of the system used in the experiment has taken the form of a computer-controlled belt/wrist band/ankle band with four embedded vibrotactile transducers: each of the small transducers are evenly spaced around the circumference of the body area (waist, wrist or ankle) and mapped to spatial audio played through a pair of headphones. The audio cues used in this experiment were created using the AM:3D [7] audio engine and were placed on a plane around the user's head at the height of the ears to avoid problems related to elevation perception. The sounds were located every 90° starting from the nose.

There were three conditions in this experiment:

Waist – four transducers are placed at cardinal points (north, south, east and west) around the waist of the participant, the waist was chosen because it has been identified as an effective body location for tactile perception and studied extensively by researchers [5, 6].

Ankle – four transducers are placed at cardinal points around the ankle of the participant. The ankle was chosen because it is an anatomical reference point with enough surface area to support four transducers and suggested by van Erp in his work on tactile navigation displays [6].

Wrist – four transducers are placed at cardinal points around the wrist on the non-dominant arm of the participant. The wrist was chosen because it is a point of mobility as suggested by Cholewiak [5] and has enough surface area to support 4 transducers at cardinal points.

The main hypothesis was that participants will be able to recognize equivalent spatial locations in an audio soundscape when given a body location and *vice versa*.

18 participants were presented with an audio or tactile cue from one of the four locations and then asked to select the equivalent cue from the choices given. For example, when the participant was presented with a *north* tactile cue on the waist, the choices presented were four different 3D audio samples. Participants had to pick the sound they believed to match the tactile version via multiple choice checkboxes.

The experimental method used a within groups design where each participant performed tasks related to all three conditions (waist, wrist, and ankle). Participants were presented with a tutorial to introduce them to the experiment and then completed 24 tasks in random order using an online system. The online system recorded one main dependent variable: the correctness of each answer. The independent variable was the different audio/tactile versions of the tactile/audio parameter.

Results

The average number of errors for the three tactile body location conditions is shown in Figure 2.

To test the hypothesis, first the significance of the effects of each tactile body location condition was investigated. The statistical analysis used here is a standard two-tailed ANOVA analysis, based on the critical values of the F distribution, with alpha=0.05.

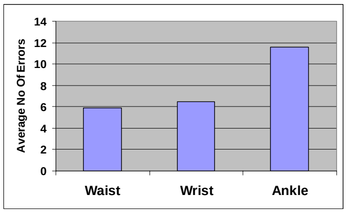

Figure 2. Average number of errors per body location.

Errors

There are significant differences in the error data between tactile body location conditions ($F=19.61 > F(2, 51) = 5.2$). Tukey's pairwise analysis showed that the average number of errors for the ankle was significantly greater than the number of errors in the waist and wrist. There were no other pairwise differences.

Qualitative Data

Participants were presented with examples of the three different tactile body locations and were asked which version that they felt was most comfortable and most easy to match with the audio equivalent. The majority of participants (66%) found the wrist to be comfortable and easiest to match with the 3D audio soundscape. No participants reported the ankle to be easy to match or particularly comfortable.

Discussion

Results show that participants are able to map the presented 3D audio positions to tactile body positions on the waist and wrist most effectively and that there are significantly more errors made when using the ankle. Although there is no significant difference between the waist and the wrist, participants indicated that they preferred the wrist.

MOBILE EXPERIMENT

Users of mobile devices are often in motion when they use their devices (e.g. receiving calls, sending text messages, etc.). Interfaces must be designed to work well under these circumstances too, not just when the user is stationary.

Given the promising results of the stationary spatial location experiment, we conducted the same experiment again in a mobile situation to see if motion affects the results. There are many ways in which motion could affect perception of crossmodal output: mobile environments tend to change frequently, the user's main attention may be on safety whilst crossing a road instead of the mobile device, a user can become physically tired, and during natural motion such as walking, a user's hands are likely to be moving.

The setup of this experiment was identical to the previous one in every respect except that a different set of participants were used and this time participants were asked to walk on a treadmill during the experiment as opposed to sitting in a chair.

Figure 3. Mobile experiment setup.

This mobile experiment used a treadmill set up in a usability lab to simulate mobility because the tactile actuators used were not wireless and were controlled from a PC and therefore inappropriate for use in a real mobile environment. Furthermore, using a treadmill permitted the experimenter to set a standard speed for all participants (in this case, all walked at a speed of 6km per hour).

The main hypothesis was that being mobile will increase errors produced during spatial location identification and matching between modalities as compared to being stationary.

Results

The average number of errors for the three tactile body location conditions is shown in Figure 4.

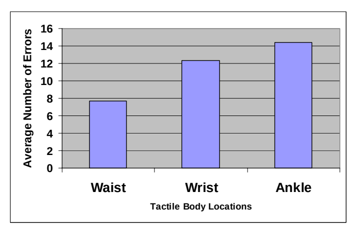

Figure 4. Average number of errors per body location.

As before, the significance of the effects of each tactile body location condition whilst mobile was investigated. Once again, the statistical analysis used here is a standard two-tailed ANOVA analysis.

Errors

There are significant differences in the error data between tactile body location conditions ($F=23.451 > F(2, 51)$). Tukey's pairwise analysis showed that the average number of errors for the ankle and wrist was significantly greater than the number of errors in the waist.

In order to establish whether there is a significant difference in the data between the stationary experiment and the mobile experiment a 2 factor ANOVA was applied using the three conditions of body location and stationary/mobile as the two factors.

The average errors for each experiment are shown below. A 2 factor ANOVA showed a significant difference between the results of mobile and stationary experiments when using the wrist and the ankle but no significant difference when using the waist.

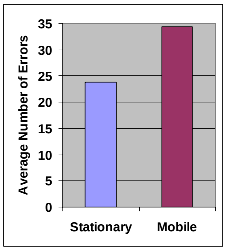

Figure 5. Average no. of errors in static and mobile conditions.

Qualitative Data

Participants were presented with examples of the three different tactile body locations and were asked which version that they felt was most comfortable and most easy to match with the audio equivalent. This time the majority of participants (83%) chose the waist.

Discussion

Results show that participants are able to map the presented 3D audio positions to tactile body positions on the waist most effectively when mobile and that there are significantly more errors made when using the ankle or wrist. Unlike the previous experiment, a greater number of participants preferred the waist to the wrist and ankle. However significantly more participants still preferred the wrist to the ankle.

The reason that the wrist performed worse in the mobile experiment compared to the static experiment could be because motion naturally changes the orientation of the wrist as the arm swings and therefore it is more difficult to match locations when they are moving constantly. For example, if we take a clock face analogy, an actuator is placed on the left hand side of the wrist to represent 0900 but as the wrist rotates during movement the actuator is no longer at the 0900 position.

ORIENTATION

In order to establish whether the natural rotation of the wrist whilst mobile confuse the interpretation of tactile cues, a further condition was tested where the arm was placed in a splint so that the wrist was unable to rotate. The experiment was otherwise the same as the mobile study.

The results show that a mobile user with a splinted wrist produces 42% fewer errors than with an unconstrained wrist. Overall, when the wrist is splinted, an ANOVA shows that there is no significant difference between the results of the mobile and static wrist condition with the wrist producing 71% correct crossmodal matches with the audio cues. This suggests that wrist rotation does cause problems, and if spatial location is to be used as a crossmodal parameter such locations would have to be avoided.

Figure 6. Wrist splint used to prevent rotation.

CONCLUSIONS

These experiments have established that it is possible for users to perceive spatial location as equivalent in both the auditory and tactile domains. Furthermore, it has been confirmed that the use of the waist as a tactile body location produces significantly better results than using the wrist or ankle. When using crossmodal spatial locations in mobile displays our experiments have shown that the wrist performs badly due to the natural rotation that takes place in motion so it is best to use the waist which, in this case, produces 76% accuracy whilst stationary and 72% accuracy whilst mobile.

Our efforts have focused on navigation as an application for displays incorporating crossmodal spatial location. However, these techniques can be utilized to encode many different types of information. Once a larger set of crossmodal parameters, including spatial location, has been established, it will be possible to include crossmodal icons in various mobile applications which allow for varying physical and social environments within which such devices are used.

REFERENCES

1. Blattner, R.M. “Earcons and icons: their structure and common design principles,” Human Computer Interaction 4(1), pp. 11-44, 1989.
2. Brewster, S.A. and Brown, L.M. “Tactons: structured tactile messages for non-visual information display”, Proc. AUI Conference, ACS, pp. 15-23, 2004.
3. Hoggan, E. and Brewster, S. “Crossmodal icons for information display”, Proceedings of CHI 2006, Canada, vol. II, pp. 857-862, 2006
4. Lewkowicz, D.J. “The development of intersensory temporal perception: an epigenetic systems/limitations view”, Psychological Bulletin, 126, pp. 281–308, 2000
5. Cholewiak, R. W., and Craig, J. C. “Vibrotactile pattern recognition and discrimination at several body sites”, Percept. Psychophys. 35, pp. 503–514, 1984
6. van Erp, J.B.F. “Tactile Navigation Display”, Lecture Notes in Computer Science, Vol. 2058, pp. 165, 2001
7. AM:3D Positional Audio, http://www.am3d.com